Sonar target enhancement by shrinkage of incoherent wavelet coefficients

Alan J. Hunter[a] and Robbert van Vossen
TNO (Netherlands Organisation for Applied Scientific Research), Oude Waalsdorperweg 63, 2597 AK The Hague, The Netherlands

(Received 13 May 2013; revised 23 October 2013; accepted 26 November 2013)

Background reverberation can obscure useful features of the target echo response in broadband low-frequency sonar images, adversely affecting detection and classification performance. This paper describes a resolution and phase-preserving means of separating the target response from the background reverberation noise using a coherence-based wavelet shrinkage method proposed recently for de-noising magnetic resonance images. The algorithm weights the image wavelet coefficients in proportion to their coherence between different looks under the assumption that the target response is more coherent than the background. The algorithm is demonstrated successfully on experimental synthetic aperture sonar data from a broadband low-frequency sonar developed for buried object detection. © *2014 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4845255]

PACS number(s): 43.60.Lq, 43.60.Hj [PJL] Pages: 262–268

I. INTRODUCTION

Broadband low-frequency sonar is a promising technology for the robust detection and classification of objects that are proud of or buried beneath the seafloor. While conventional high-frequency imaging sonar relies on geometrical features (such as the highlights and shadows) to detect and classify objects, broadband low-frequency sonar can be used to exploit structural resonances of the objects.[1,2] The classification of contacts in broadband low-frequency sonar imagery thus depends on a capability to reliably extract the resonant echo response. In practice, the echo response is corrupted by background reverberation, which can adversely affect this capability. This is a problem for low-frequency sediment-penetrating sonars, e.g., the Netherlands Organisation for Applied Scientific Research (TNO) MUD,[3] SSAM,[4] and BOSS[5] systems, where the seafloor sediment and/or multipath reverberation can be strong while useful features of the echo response can be comparatively weak.

Synthetic aperture sonar (SAS) processing provides increased image resolution by exploiting the inter-ping phase coherence of the echo data. Furthermore, it enhances the signal-to-noise ratio for coherent targets relative to the incoherent background reverberation.[6] Use of a directional source and/or receiver array can further mitigate the multipath reverberation in shallow water.[7] However, even with these mitigation measures in place, the residual reverberation can be problematic.

A target's multi-aspect acoustic color and the echo response from which it is derived are potentially useful representations for classifying targets in broadband low-frequency sonar data.[8] The response of a target of interest can be isolated from the background reverberation and other nearby targets in the SAS image by applying a window followed by a reversal of the SAS imaging procedure.[8,9] However, there is a trade-off for the choice of window size, which is driven primarily by the need to suppress the reverberation noise; a smaller window achieves greater suppression of the reverberation, but at the cost of reduced spatial frequency resolution and a risk of losing important target features. Furthermore, any reverberation remaining within the window will contribute to a degraded target response.

Ideally, the target can be isolated from the background reverberation without sacrificing resolution or losing target information. Recently, a novel coherence-based wavelet shrinkage method was proposed for de-noising magnetic resonance images that is capable of achieving this.[10,11] Here, we adapt this method for the purpose of separating coherent targets from the incoherent background reverberation in broadband low-frequency SAS images. By removing the reverberation in this way, a window is only required to isolate the target of interest from other targets and, thus, a much larger window can be employed reducing the risk that important information is lost. We demonstrate application of the method on experimental data from the TNO MUD sediment-penetrating sonar.

II. ALGORITHM DESCRIPTION

A variety of image de-noising algorithms has been developed based on wavelet shrinkage using the magnitudes of the wavelet coefficients; the SUREshrink algorithm[12] is a popular example. However, in sonar imagery the magnitude of the echoes is not always the best criterion for separating targets from the background reverberation noise (consider, for example, a weakly scattering target on a highly reverberant seafloor). A key characteristic that does distinguish targets from background reverberation is echo structure, and this can be quantified by the coherence.

The proposed method uses a coherence metric to determine the similarity of wavelet coefficients between independent looks, i.e., different images of the same scene with statistically independent noise realizations. The wavelet coefficients that have high coherence between looks are assumed to correspond to the reverberation-free measurements of the targets (i.e., structured echoes), whereas the coefficients with

[a]Author to whom correspondence should be addressed. Electronic mail: alan.hunter@tno.nl

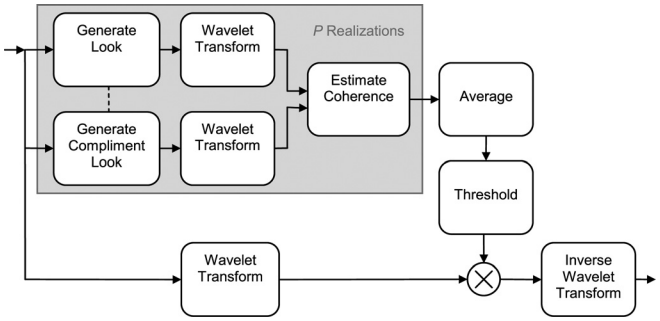

FIG. 1. Overview of the wavelet shrinkage algorithm.

low coherence are assumed to correspond to the reverberation. Under these assumptions, the coefficients with low coherence are attenuated to produce a de-noised image. The adaptation of this method to the sonar application is described in Secs. II A–II C and is summarized in Fig. 1.

A. Generation of complementary looks

In the context of target detection and classification for broadband low-frequency sonar, the major source of interference is from the seafloor sediment and multipath reverberation. Here, we assume that the reverberation cannot be resolved by the system and, therefore, behaves as speckle, i.e., it has low coherence over the spatial and temporal frequencies.[13] Comparatively, targets are assumed to be much more coherent. These are the characteristics of the data that are exploited to generate statistically independent looks.

A broadband low-frequency synthetic aperture sonar collects samples over a wide range of temporal and spatial frequencies due to the use of a broadband pulse and wide beams, respectively. Therefore, the SAS echo data have broad support in the wavenumber domain. Complementary looks can thus be generated by partitioning this space into two sets. The partitioning of wavenumber space must be made so that this relationship is preserved in order to exploit the assumed coherence of the target and relative incoherence of the background. One means of achieving this is to apply a two-dimensional (2-D) grid to the space and assign blocks to either one or the other set; this is illustrated in Fig. 2 for a regular assignment between sets. It is necessary to choose a grid such that the speckle is correlated only on a scale that is much smaller than the block size and the target is correlated on a scale that is bigger. In this way, the targets are likely to be persistent between the blocks whereas the speckle is not. The derivation of an optimal grid will depend on the assumed scattering properties of the targets and reverberation, and the system parameters; this is an interesting subject for future research, but is not considered here.

Often in practice, the data are sampled at close to the Nyquist rate, or perhaps even undersampled, and a regular separation of blocks between sets (as illustrated in Fig. 2) can result in undersampling and the introduction of artifacts. In these cases, it is advantageous to partition the space using a randomized assignment between sets, thus breaking the regularity and spreading the energy of the undersampling artifacts rather than concentrating them into localized grating lobes. Here, we use a randomized partitioning scheme where

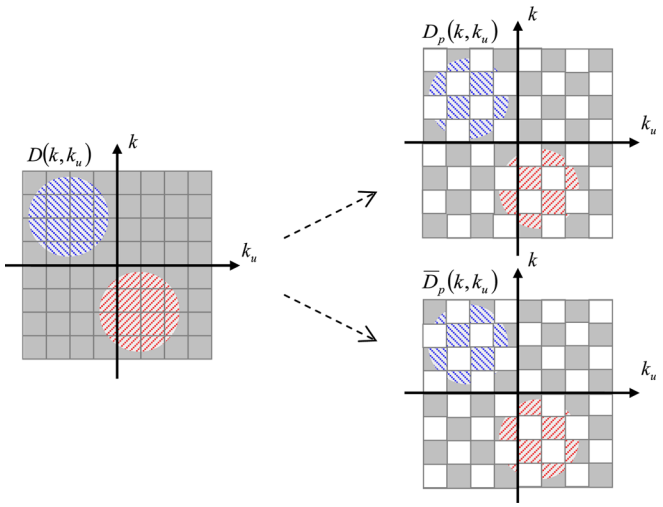

FIG. 2. (Color online) Generation of complementary looks by a regular partitioning of the wavenumber space where alternate blocks are assigned into two sets. The hashed areas represent regions of the space corresponding to target responses that are correlated over multiple blocks.

blocks are assigned to either of the two complementary sets (looks) at random, i.e.,

$$D_p(k,k_u) = D(k,\,k_u)w_p(k,\,k_u), \tag{1}$$

$$\tilde{D}_p(k,\,k_u) = D(k,\,k_u)\big[1 - w_p(k,\,k_u)\big], \tag{2}$$

where $D(k,k_u) = \mathrm{FT}\{d(t,u)\}$ is the 2-D Fourier transform of the echo data $d(t,u)$, t and u are time and the along-track positions, respectively, and k and k_u are their respective wavenumbers;

$$w_p(k,\,k_u) = a_p\left[\mathrm{round}\left(\frac{k-k_0}{\Delta k}\right),\,\mathrm{round}\left(\frac{k_u-k_{u,0}}{\Delta k_u}\right)\right] \tag{3}$$

is a window function that implements the random block assignment over a grid with centers at the wavenumbers $(k_m,k_{u,n}) = (k_0 + m\Delta k, k_{u,0} + n\Delta k_u)$, where $a_p[m,n]$ is a 2-D pseudo-random noise sequence drawn from a binary distribution, round(x) is the nearest integer to x, and the subscript p denotes the pth sequence realization. The SAS images for a pair of complementary looks, $i_p(x,y)$ and $\tilde{i}_p(x,y)$, can then be generated from the data, i.e.,

$$i_p(x,y) - \mathrm{IM}\{\mathrm{FT}^{-1}\{D_p(k_u,k)\}\}, \tag{4}$$

$$\tilde{i}_p(x,y) = \mathrm{IM}\{\mathrm{FT}^{-1}\{\tilde{D}_p(k_u,\,k)\}\}, \tag{5}$$

where $\mathrm{IM}\{\cdot\} : t,\,u \rightarrow x,\,y$ is the SAS imaging operator, and x and y are the along-track and across-track image coordinates, respectively.

B. Coherence of wavelet coefficients

The wavelet domain is good for yielding sparse image representations, meaning that the important information tends to be contained within only a few coefficients whereas noise is distributed more evenly. It is, therefore, a good

domain for separating targets from noise and has been used for resolution-preserving de-noising in imaging applications, including sonar.[14]

The images for the complementary looks can be transformed to the wavelet domain by the discrete wavelet transform, i.e.,

$$C_p(x', y', \mathbf{s}) = \mathrm{WT}\{i_p(x, y)\}, \tag{6}$$

$$\tilde{C}_p(x', y', \mathbf{s}) = \mathrm{WT}\{\tilde{i}_p(x, y)\}, \tag{7}$$

where the operator $\mathrm{WT}\{\cdot\} : x,\ y \rightarrow x',\ y',\ \mathbf{s}$ denotes the wavelet transform and $\mathbf{s} = (s_x, s_y, s_{xy})$ is the wavelet scale.

A thorough investigation into the optimal wavelet type was not conducted. However, a wavelet with symmetric/anti-symmetric real and imaginary components should be chosen to preserve the phase information in the image;[15] this can be important for subsequent signal processing. The coefficients corresponding to targets or noise are distinguished by considering a measure of their similarity between the looks; coefficients that are similar are designated as belonging to a potential target, whereas those that are dissimilar are designated as noise. The similarity is quantified here using the coherence, but other metrics can also be used.[10,11] The coherence between looks can be estimated using

$$\rho_p(x', y', \mathbf{s}) = \frac{\sum_{m,n=-(N-1)/2}^{(N-1)/2} C_p(x' + m\Delta x,\ y' + n\Delta y, \mathbf{s})\bar{C}_p^*(x' + m\Delta x,\ y' + n\Delta y, \mathbf{s})}{\sqrt{\sum_{m,n=-(N-1)/2}^{(N-1)/2} |C_p(x' + m\Delta x,\ y' + n\Delta y, \mathbf{s})|^2}\sqrt{\sum_{m,n=-(N-1)/2}^{(N-1)/2} |\bar{C}_p^*(x' + m\Delta x,\ y' + n\Delta y, \mathbf{s})|^2}} \tag{8}$$

for an $N \times N$ averaging window, where Δx, Δy are the pixel dimensions. The variance of the coherence estimate can be reduced by averaging over a larger window (i.e., larger N) at the cost of reduced spatial resolution. The variance can also be reduced by averaging over a number, P, of different look realizations, i.e.,

$$\rho(x', y', s) \approx \frac{1}{P}\sum_{p=1}^{P} \rho_p(x', y', \mathbf{s}), \tag{9}$$

where, for example, the different looks are obtained by repartitioning the data with different realizations of the pseudo-random noise sequence.

C. Shrinkage of wavelet coefficients

Having determined the coherence of the wavelet coefficients, a weighting scheme is chosen to attenuate those coefficients that are deemed to be incoherent and, therefore, designated as noise. The derivation of an optimal weighting scheme will depend on the assumed properties of the targets, background reverberation, and the system parameters; this is an interesting subject for future research. However, to demonstrate the principle we define here a simple *ad hoc* scheme with a user-defined soft threshold, i.e., with weights

$$A(x', y', \mathbf{s}) = \begin{cases} 1, & \rho(x', y', \mathbf{s}) > \rho_{\max} \\ 0, & \rho(x', y', \mathbf{s}) < \rho_{\min} \\ \dfrac{\rho(x', y', \mathbf{s}) - \rho_{\min}}{\rho_{\max} - \rho_{\min}}, & \text{otherwise,} \end{cases} \tag{10}$$

where coefficients with coherence less than $\rho_{\min}$ are attenuated completely and coefficients greater than $\rho_{\max}$ are retained completely, with a linear graduation between these thresholds.

Using these coherence-based weights, the wavelet coefficients of the full image are adjusted and then inverse wavelet transformed to yield the image containing only the coherent targets, i.e.,

$$i'(x, y) = \mathrm{WT}^{-1}\{A(x', y', \mathbf{s})\mathrm{WT}\{i(x, y)\}\}, \tag{11}$$

where

$$i(x, y) = \mathrm{IM}\{\mathrm{FT}^{-1}\{D(k_u, k)\}\} \tag{12}$$

is the full SAS image. Note that an image of the background reverberation can be obtained by subtracting the coherent target image from the original image, i.e., $i(x, y) - i'(x, y)$.

III. RESULTS FROM THE MUD SONAR

The MUD system is a low-frequency broadband synthetic aperture sonar developed by TNO for experimentation on buried object detection.[3] The system wet end is shown in Fig. 3. It is comprised of an exchangeable acoustic projector, two 16-element hydrophone arrays, and navigation sensors (real-time kinematic global positioning system and a photonic inertial navigation system). These components are mounted on a frame with an adjustable tilt angle that is adapted for operation from diver support vessels of the Royal Netherlands Navy.

Using a 4–9 kHz projector and a sidescan imaging geometry, the system is capable of detecting objects completely buried in mud at ranges up to 70 m in very shallow water (approximately 10–15 m water depth and 0.5 m of

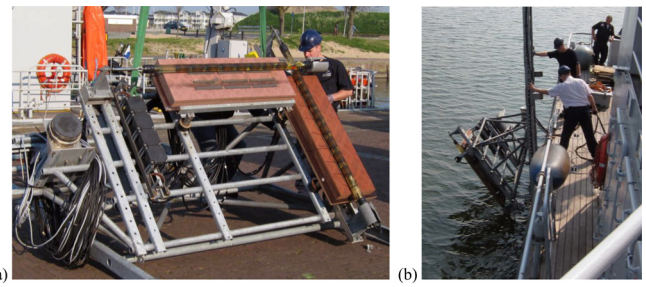

FIG. 3. (Color online) MUD low-frequency broadband synthetic aperture sonar: (a) wet-end, (b) deployment.

mud). Figure 4(a) shows a SAS image acquired by the system during a 2011 sea trial in the Haringvliet.[8] A test garden consisting of a variety of different objects was deployed for the trial, including mine-like and non-mine-like objects, unexploded ordinance (UXO), and calibration targets. These objects were actively buried by divers six months prior to the trial. A high-frequency sonar survey was conducted immediately before the trial using a REMUS autonomous underwater vehicle and this confirmed that the objects were indeed buried (i.e., they were not detected during the REMUS survey). The contacts corresponding to the deployed objects were identified using ground-truth measurements and these are indicated by arrows in the SAS image of Fig. 4(a). This demonstrates the capability of the system to detect the buried objects. However, in addition to the objects of interest, the system also detects regions of dense clutter, an example of which has been indicated within the boxed area of Fig. 4(a). In some cases, the high clutter density prevents the confident matching of target contacts with the ground-truth. It is clear from this example that robust classification is required to discriminate between targets and clutter in order to achieve an operational capacity.

Isolation of the contact responses from the background reverberation is a crucial preprocessing step in the classification procedure. The proposed wavelet shrinkage technique is demonstrated here for this purpose. The technique was applied to the SAS image of Fig. 4(a) using the following parameters: the data support in wavenumber space was partitioned into a grid of 32×32 blocks to generate the complementary looks; a window of $N = 5$ pixels was used at each wavelet scale to estimate the coherence (image pixel size of $\Delta x = 0.2\,\mathrm{m}$ and $\Delta y = 0.1\,\mathrm{m}$); the coherence was averaged over $P = 100$ look realizations; a Morlet wavelet[16] was used in the wavelet transform; and a soft threshold of

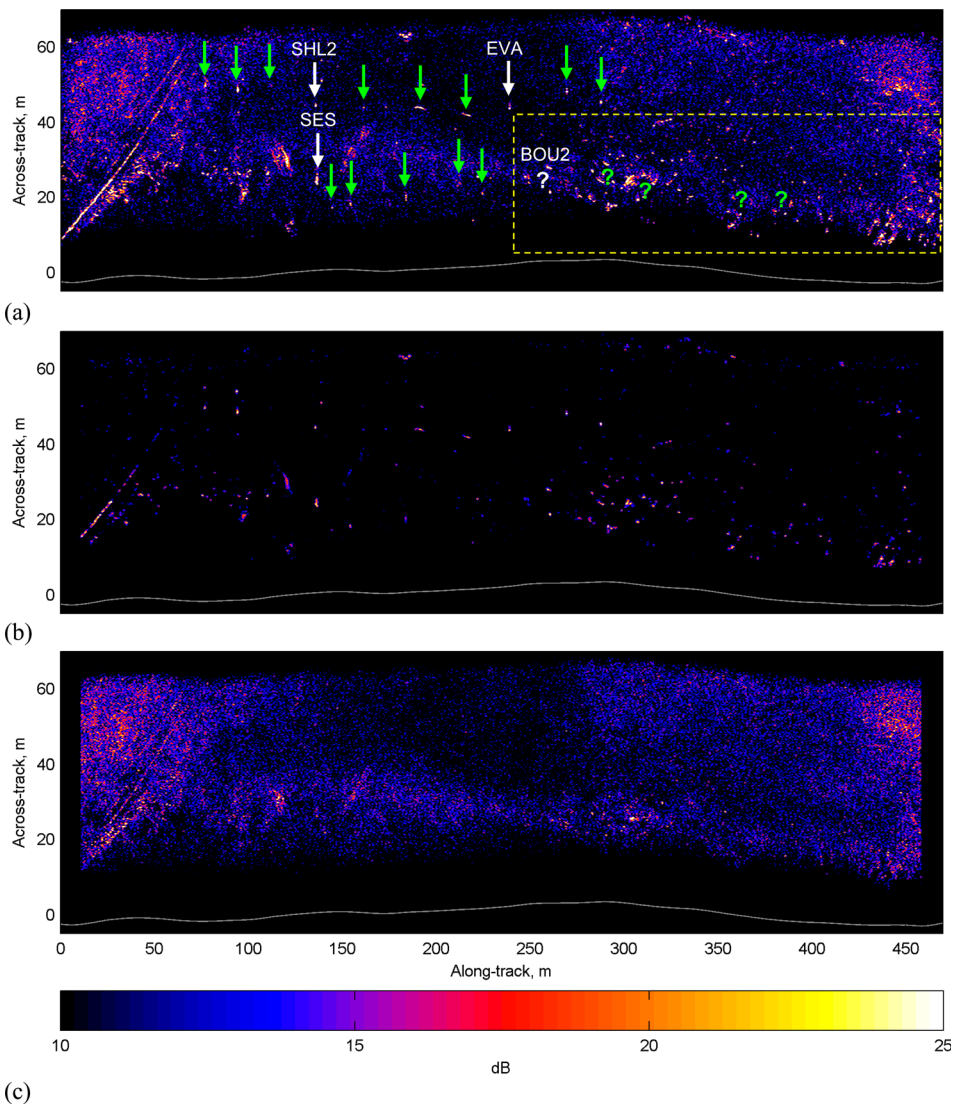

FIG. 4. (Color online) Run 325 from the MUD-2011 trials. (a) SAS image with contacts of the deployed targets indicated by arrows and an area of dense clutter indicated within the boxed region; for this run, the contacts corresponding to deployed objects could not be identified with confidence in this region and the ground truth locations are indicated by the question marks; (b) coherent contacts; and (c) background noise separated using the wavelet shrinkage method.

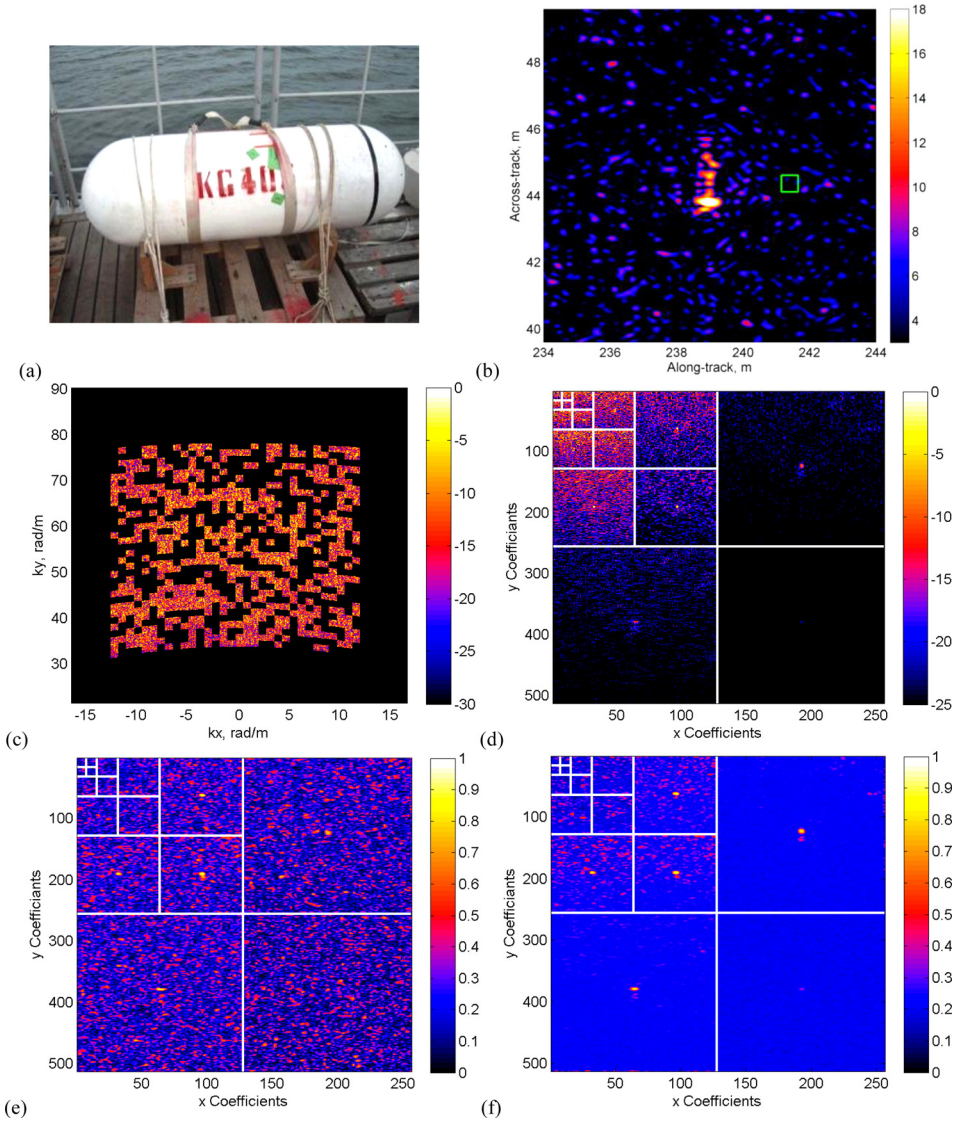

FIG. 5. (Color online) Target/background separation for the CMRE EVA cylinder target: (a) photograph of the EVA cylinder, (b) SAS image snippet (the green square denotes the ground-truth location), (c) wavenumber spectrum of a random look realization, (d) wavelet decomposition of the look image, (e) wavelet coherence between a pair of complementary looks, and (f) average wavelet coherence over 100 look realizations.

$\rho_{\min} = 0.4$ and $\rho_{\max} = 0.5$ was chosen to weight the wavelet coefficients. The separated contacts and background are shown in Figs. 4(b) and 4(c), respectively. Qualitatively, it can be observed that the contacts have been well isolated and the resolution has been preserved.

The procedure is shown in more detail in Fig. 5 for a specific target: the EVA cylinder[17] from NATO's Centre for Maritime Research and Experimentation (CMRE). A photograph of the EVA cylinder is shown in Fig. 5(a). The cylinder is 1.5 m long and 0.5 m in diameter with hemispherical end-caps, it is constructed from a thin shell of composite material, and filled with water except for one end-cap which is solid epoxy resin. Figure 5(b) shows the SAS image snippet. A single realization of the randomly partitioned wavenumber space used to generate a look is shown in Fig. 5(c), the wavelet transform of the look image is shown in Fig. 5(d), and its coherence with respect to the complementary look image is shown in Fig. 5(e). The coherence of the wavelet coefficients between the complementary looks is very noisy for a single realization, but this is reduced significantly by averaging the coherence over multiple complementary look realizations, as shown in Fig. 5(f).

The SAS image snippet for the isolated response of the EVA cylinder is shown in Fig. 6(a). The corresponding echo data (obtained by reversal of the SAS imaging operator) are shown in Fig. 6(c) and the multi-aspect acoustic color (i.e., the aspect-frequency representation of the echo data) is shown in Fig. 6(e). For comparison, Figs. 6(b), 6(d), and 6(f) show, respectively, the SAS snippet, echo data, and multi-aspect acoustic color, where the target isolation has been achieved by application of a window of dimensions 2.5×5 m with a 0.5 m transition border. In this case, the window is too large and some of the background is retained, thus corrupting the target response. Conversely, when the window is too small important details can be lost.

A. Quantitative analysis

The procedure was repeated for a variety of targets over different runs to gain a quantitative estimate of the average improvement offered by the method. The targets considered were: (i) the EVA cylinder, (ii) a 155 mm shell (UXO), (iii) a boulder of approximately 40 cm diameter, and (iv) SESAME, a pressurized container containing underwater acoustic logging equipment. These targets are indicated in

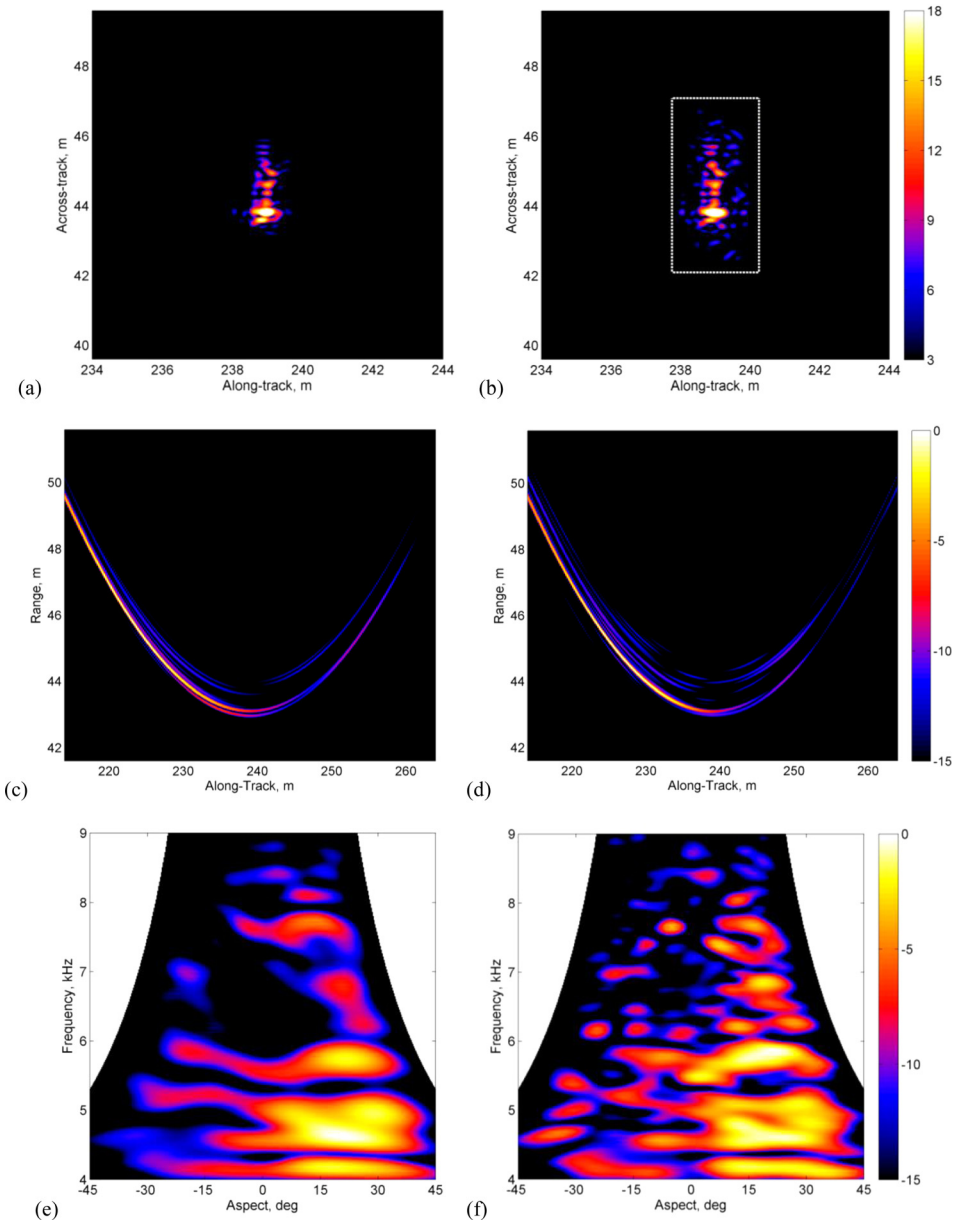

FIG. 6. (Color online) Acoustic signatures of the EVA target obtained using: (a),(c),(e) the wavelet shrinkage method with a soft threshold for the coherence between 0.4 and 0.5; and (b),(d),(f) a spatial window of dimensions 2.5×5 m. (a),(b) shows the SAS image snippet, (c),(d) the echo data, and (e),(f) the multi-aspect acoustic color.

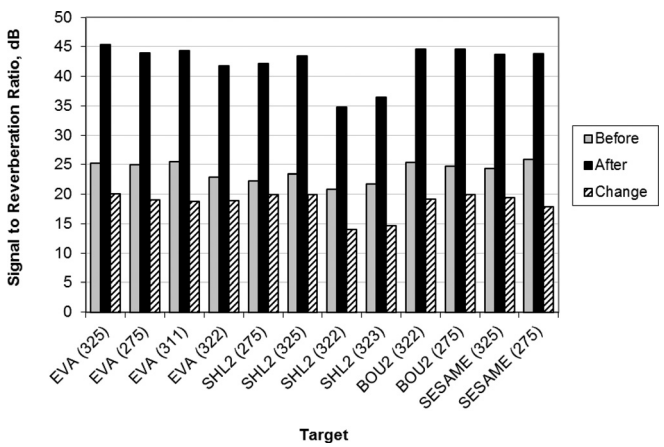

FIG. 7. Comparison of estimated signal-to-reverberation ratio before and after target enhancement by incoherent wavelet shrinkage. The labels EVA, SHL2, BOU2, and SES denote the cylinder, shell, boulder, and SESAME targets, respectively, and the numbers in parentheses indicate the run in which the target was observed.

the SAS image of Fig. 4 by the identifiers EVA, SHL2, BOU2, and SES, respectively. The targets were observed over several different runs at varying ranges and orientations relative to the sonar, where the runs selected here are denoted by the numbers 275, 311, 322, 323, and 325 (run 325 is shown in Fig. 4).

The improvement is quantified using an estimate of the signal-to-reverberation ratio (SRR) given by

$$\mathrm{SRR} = \max(|i(x,y)|^2)/\mathrm{median}(|i(x,y)|^2) \tag{13}$$

over an image snippet of dimensions 10×10 m centered on the target. The results are tabulated in Fig. 7 and show an average SRR improvement between 15 and 20 dB.

IV. CONCLUSIONS

A wavelet shrinkage algorithm has been presented for the enhancement of the acoustic response of coherent targets

in low-frequency synthetic aperture sonar imagery. The algorithm has been demonstrated on data from the TNO MUD low-frequency sonar and shown to suppress the background seafloor reverberation effectively, achieving an average increase in SRR of between 15 and 20 dB while preserving image resolution and phase. Further work is required to investigate the optimal generation of complementary looks based on prior knowledge of the reverberation and target coherence and optimal thresholding schemes.

ACKNOWLEDGMENTS

This work was partially funded by the U.S. Department of Defense Strategic Environmental Research and Development Program (SERDP), and the authors thank Dr. Herb Nelson of SERDP for his support. The authors also acknowledge the Royal Netherlands Navy for funding and operational support for the MUD sea trials and the NATO Centre for Maritime Research and Experimentation (CMRE) for providing the EVA cylinder.

[1]J. A. Bucaro, B. H. Houston, M. Saniga, L. R. Dragonette, T. Yoder, S. Dey, L. Kraus, and L. Carin, "Broadband acoustic scattering measurements of underwater unexploded ordinance," J. Acoust. Soc. Am. **123**(2), 738–746 (2008).

[2]K. L. Williams, S. G. Kargl, E. I. Thorsos, D. S. Burnett, J. L. Lopez, M. Zampolli, and P. L. Marston, "Acoustic scattering from a solid aluminum cylinder in contact with a sand sediment: Measurements, modeling, and interpretation," J. Acoust. Soc. Am. **127**(6), 3356–3371 (2010).

[3]A. L. D. Beckers, R. van Vossen, and G. Vlaming, "Low-frequency synthetic aperture sonar for detecting explosives in harbours," Sea Technol. **53**(3), 15–18 (2012).

[4]D. D. Sternlicht, J. E. Fernandez, R. Holtzapple, D. P. Kucik, T. C. Montgomery, and C. M. Loeffler, "Advanced sonar technologies for autonomous mine countermeasures," in *Proceedings MTS/IEEE Oceans 2011*, Waikoloa, HI (2011), pp. 1–5.

[5]S. G. Schock, J. Wulf, G. Quentin, and J. Sara, "Synthetic aperture processing of buried object scanning sonar data," in *Proceedings MTS/IEEE Oceans2005*, Washington, DC (2005), Vol. 3, pp. 2236–2241.

[6]B. J. Davis, P. T. Gough, and B. R. Hunt, "Modeling surface multipath effects in synthetic aperture sonar," IEEE J. Ocean. Eng. **34**(2), 239–249 (2009).

[7]M. Pinto, A. Bellettini, L. S. Wang, P. Munk, V. Myers, and L. Pautet, "A new synthetic aperture sonar design with multipath mitigation," in *Proceedings AIP High Frequency Ocean Acoustics Conference*, La Jolla, CA (2004), pp. 489–496.

[8]A. J. Hunter, R. van Vossen, B. A. J. Quesson, M. E. G. D. Colin, M. Zampolli, and A. L. D. Beckers, "Low frequency synthetic aperture sonar for detecting and classifying buried objects," in *Proceedings 11th European Conference on Underwater Acoustics*, Edinburgh, UK (2012), pp. 1–8.

[9]P. Marston, T. Marston, and K. L. Williams, "Scattering resonances, filtering with reversible SAS processing, and applications of quantitative ray theory," in *Proceedings MTS/IEEE Oceans2010*, Seattle, WA (2010), pp. 1–9.

[10]O. Tischenko, C. Hoeschen, and E. Buhr, "An artefact-free structure-saving noise reduction using the correlation between two images for threshold determination in the wavelet domain," Proc. SPIE **5747**, 1066–1075 (2005).

[11]R. A. Borsdorf, R. Raupach, T. Flohr, and J. Hornegger, "Wavelet based noise reduction in CT images using correlation analysis," IEEE Trans. Med. Imaging **27**(12), 1685–1703 (2008).

[12]D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," J. Am. Stat. Assoc. **90**(432), 1200–1224 (1995).

[13]J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications* (Roberts and Co., Greenwood Village, CO, 2006), pp. 1–384.

[14]B. Ergen, "Signal and image denoising using wavelet transform," in *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology* (InTech, Rijeka, Croatia, 2012), Chap. 21, pp. 495–514; available at www.intechopen.com (Last viewed April 2012).

[15]P. Kovesi, "Phase preserving denoising of images," in *Proceedings APRS Digital Image Computing: Techniques and Applications*, Perth, Australia (1999), pp. 1–6.

[16]J. Mortlet, G. Arens, E. Fourgeau, and D. Giard, "Wave propagation and sampling theory—Part I: Complex signal and scattering multilayered media," Geophysics **47**(2), 203–221 (1982).

[17]A. Tesei, M. Zampolli, and G. Canepa, "At-sea measurements of acoustic elastic scattering by a 1.5 m long cylinder made of composite materials," in *Proceedings 2nd International Conference on Underwater Acoustic Measurements*, Heraklion, Greece (2007), pp. 1–6.